

# Exact Online String Matching Bibliography

Simone Faro

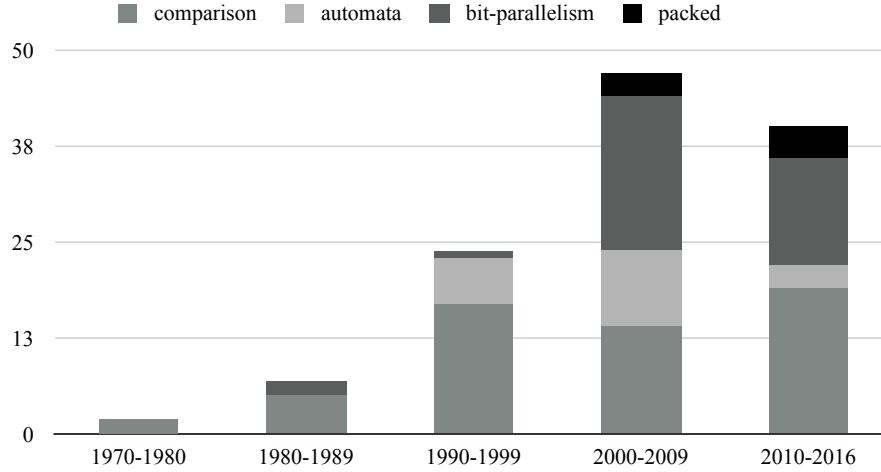
Università degli Studi di Catania, Dipartimento di Matematica e Informatica  
Viale Andrea Doria 6, I-95125, Catania, Italy  
`faro@dmf.unict.it`

**Abstract.** In this short note we present a comprehensive bibliography for the *online exact string matching problem*. The problem consists in finding *all* occurrences of a given pattern in a text. It is an extensively studied problem in computer science, mainly due to its direct applications to such diverse areas as text, image and signal processing, speech analysis and recognition, data compression, information retrieval, computational biology and chemistry. Since 1970 more than 120 string matching algorithms have been proposed. In this note we present a comprehensive list of (almost) all string matching algorithms. The list is updated to May 2016.

## 1 Introduction

Given a text  $t$  of length  $n$  and a pattern  $p$  of length  $m$  over some alphabet  $\Sigma$  of size  $\sigma$ , the *string matching problem* consists in finding *all* occurrences of the pattern  $p$  in the text  $t$ . String matching algorithms are also basic components used in implementations of practical softwares existing under most operating systems. Moreover, they emphasize programming methods that serve as paradigms in other fields of computer science. Finally they also play an important role in theoretical computer science by providing challenging problems. Applications require two kinds of solutions depending on which string, the pattern or the text, is given first. Algorithms based on the use of automata or combinatorial properties of strings are commonly implemented to preprocess the pattern and solve the first kind of problem. This kind of problem is generally referred as *online* string matching. The notion of indexes realized by trees or automata is used instead in the second kind of problem, generally referred as *offline* string matching. In this note we are only interested in algorithms of the first kind.

Since 1970 more than 120 online string matching algorithms have been proposed. All solutions can be divided into several different classes. Here we divide all algorithms into four classes, although different classifications are possible. Specifically we distinguish algorithms which solve the problem by making use of comparisons between characters, algorithms based on packed string matching, and algorithms which make use of automata. The latter class can be further divided into two classes: algorithms which make use of deterministic automata



**Fig. 1.** Number of appeared string matching algorithms since 1970.

and algorithms based on bit-parallelism which simulate the behavior of non-deterministic automata.

The lists of string matching algorithms proposed in the following sections are in chronological ordered. Each entry of the lists includes the name of the algorithm, a small acronym used in literature to refer to the algorithm, the list of references to the paper where the algorithm appeared and a very short description of the main features of the algorithm. The Smart tool<sup>1</sup> includes the implementation of (almost) all the algorithms presented in this note. The small acronym associated to each algorithm is used in Smart to refer to the algorithm.

## List of Algorithms based on Comparison of Characters

In a computational model where the matching algorithm is restricted to read all the characters of the text one by one the optimal complexity is  $\mathcal{O}(n)$ , and was achieved the first time by the well known Morris-Pratt algorithm [80]. However, in many practical cases it is possible to avoid reading all the characters of the text achieving sub-linear performances on the average.

The optimal average  $\mathcal{O}(\frac{n \log_{\sigma} m}{m})$  time complexity [99] was reached by many comparison based algorithms. However, all algorithms with a sub-linear average behavior may have to read all the text characters in the worst case. It is interesting to note that many of those algorithms have an even worse  $\mathcal{O}(nm)$ -time complexity in the worst-case [48,52].

<sup>1</sup> Smart (String Matching Algorithms Research Tool), by Simone Faro and Thierry Lecroq, is available at <http://www.dmi.unict.it/~faro/smart/>

Int what follows we present the list of all string matching algorithms based on comparison of characters.

1. **Brute-Force** (BF) - no date  
 Appeared in [31]  
 Naïve algorithm for the string matching problem. It works in  $\mathcal{O}(nm)$ -time
2. **Morris-Pratt** (MP) - 1970  
 Appeared in [80]  
 First linear algorithm scanning the text and the pattern from left to right. It derives from (1). Preprocessing of the pattern in  $\mathcal{O}(m)$ -time and -space.
3. **Knuth-Morris-Pratt** (KMP) - 1977  
 Appeared in [74]  
 Linear algorithm. Improvement of (2).
4. **Boyer-Moore** (BM) - 1977  
 Appeared in [18]  
 Introduces the sliding window approach and the scanning of the window from right to left. It works in  $\mathcal{O}(nm)$ -time. It uses the occurrence and the good-suffix heuristics. It derives from (1)
5. **Horspool** (HOR) - 1980  
 Appeared in [65]  
 Improvement of (4). It uses only a modification if the occurrence heuristics.
6. **Galil-Seiferas** (GS) - 1981  
 Appeared in [60,61]  
 Linear algorithm using constant extra space complexity. Preprocessing in  $\mathcal{O}(m)$ -time. It performs  $5n$  text character comparisons in the worst case.
7. **Apostolico-Giancarlo** (AG) - 1986  
 Appeared in [7]  
 Variant of (4). It works in  $\mathcal{O}(n)$ -time complexity and performs  $3n/2$  comparisons in the worst case.
8. **Karp-Rabin** (KR) - 1987  
 Appeared in [73]  
 First filter algorithm using an hashing function. It works in  $\mathcal{O}(mn)$ -time complexity but has an  $\mathcal{O}(n + m)$  expected running time.
9. **Zhu-Takaoka** (ZT) - 1987  
 Appeared in [101]  
 Improvement of (5) which uses two characters for computing the occurrence heuristics.
10. **Quick-Search** (QS) - 1990  
 Appeared in [96]  
 Improvement of (5). It uses the character which follows the current window to compute the occurrence heuristics.

11. **Optimal-Mismatch** (OM) - 1990  
 Appeared in [96]  
 Variation of (10) using probabilities of characters. It compares the pattern and the text proceeding from the less frequent to the more frequent character.
12. **Maximal-Shift** (MS) - 1990  
 Appeared in [96]  
 Variation of (10) where pattern characters are scanned from the one which will lead to a larger shift to the one which will lead to a shorter shift
13. **Apostolico-Crochemore** (AC) - 1991  
 Appeared in [6]  
 Modification of (3). It performs at most  $3n/2$  characters inspections in the worst case.
14. **Two-Way** (TW) - 1991  
 Appeared in [34]  
 Linear algorithm in the worst case. Divides the pattern in two factors. It proceeds from left to right while scanning the left factor, and proceeds from right to left while scanning the right factor. It inspects at most  $2n - m$  characters.
15. **Tuned-Boyer-Moore** (TUNBM) - 1991  
 Appeared in [69,70]  
 Variant of (5) which introduces a fast loop with unrolled blind shifts.
16. **Colussi** (COL) - 1991  
 Appeared in [29]  
 Refinement of (3) using a 2-factorization of the pattern.
17. **Smith** (SMITH) - 1991  
 Appeared in [93]  
 Combination of (5) and (10) which takes the maximum of the shifts proposed by the two occurrences heuristics.
18. **Galil-Giancarlo** (GG) - 1992  
 Appeared in [59]  
 Refinement of (16) which performs at most  $4n/3$  inspections in the worst case.
19. **Raita** (RAITA) - 1992  
 Appeared in [88]  
 Modification of (5). It first compare the last character, then the first, and finally the middle character. If no mismatch occurs all other characters are compared.
20. **String-Matching on Ordered ALphabet** (SMOA) - 1992  
 Appeared in [32]  
 It uses constant extra space and achieves a linear time complexity. It inspects at most  $6n + 5$  characters.

21. **Turbo-Boyer-Moore** (TBM) - 1992  
 Appeared in [53,33]  
 Improvement of (4). It remembers characters inspected in the previous attempt. It performs at most  $2n$  character inspections.
22. **Not-So-Naive** (NSN) - 1993  
 Appeared in [62]  
 Improvement of (1). It uses constant extra space and works in  $\mathcal{O}(nm)$ -time.
23. **Reverse-Colussi** (RCOL) - 1994  
 Appeared in [30]  
 Combination of (16) and (4). It works in linear time and performs at most  $2n$  character inspections.
24. **Skip-Search** (SKIP) - 1998  
 Appeared in [28]  
 It introduces the use of buckets of positions for each character of the alphabet. It works in  $\mathcal{O}(nm)$ -time but has a linear average behavior.
25. **Alpha-Skip-Search** (ASKIP) - 1998  
 Appeared in [28]  
 Improvement of (24). It uses buckets of positions for each factor of length  $\log \sigma(m)$  of the pattern.
26. **Knuth-Morris-Pratt Skip-Search** (KMPS) - 1998  
 Appeared in [28]  
 Improvement of (24) using shift tables in (2) and (3).
27. **Berry-Ravindran** (BR) - 1999  
 Appeared in [15]  
 Combination of (10) and (9). It uses the two consecutive characters, after the current window, for computing the occurrence heuristics.
28. **Ahmed-Kaykobad-Chowdhury** (AKC) - 2003  
 Appeared in [1]  
 A variant of (7) that remembers all the suffixes of the pattern found in the text and that computes the shifts accordingly at the end of each attempt.
29. **Fast-Search** (FS) - 2003  
 Appeared in [21,22]  
 Improvement of (4). It uses the occurrence heuristics when the mismatch occurs during the first comparison.
30. **Forward-Fast-Search** (FFS) - 2004  
 Appeared in [22]  
 Combination of (29) and (10). It computes the good-suffix using information about the character following the current window of the text.
31. **Backward-Fast-Search** (BFS) - 2004  
 Appeared in [22]  
 Combination of (4) and (5). It computes the good-suffix heuristics using information about the mismatching character.

32. **Tailed-Substring** (TS) - 2004  
 Appeared in [19]  
 Improvement of (1) using a variant of the occurrence heuristics in (5). It uses constant extra space.
33. **Sheik-Sumit-Anindya-Balakrishnan-Sekar** (SSABS) - 2004  
 Appeared in [91]  
 Combination of (10) and (19).
34. **Thathoo-Virmani-Sai-Balakrishnan-Sekar** (TVSBS) - 2006  
 Appeared in [66]  
 Combination of (33) and (27).
35. **Boyer-Moore-Horspool using Probabilities** (PBMH) - 2006  
 Appeared in [84]  
 Combination of (11) and (5).
36. **Franek-Jennings-Smyth** (FJS) - 2007  
 Appeared in [54,55]  
 Combination of (3) and (10).
37. **2-Block Boyer-Moore** (2BLOCK) - 2007  
 Appeared in [97]  
 Improvement of (4) performing a constant number of inspections in the worst case.
38. **Wu-Manber for Single Pattern Matching** (HASHQ) - 2007  
 Appeared in [78]  
 Improvement of (5) using a super alphabet. It computes a fingerprint of each  $q$ -gram in the pattern using an hash function.
39. **Boyer-Moore-Horspool with  $q$ -grams** (BMHQ) - 2008  
 Appeared in [72]  
 Improvement of (5) using a super alphabet. It computes the occurrence heuristics reading  $q$ -grams in a single operation.
40. **Two Sliding Windows** (TSW) - 2008  
 Appeared in [67]  
 Improvement of (10) using two sliding windows. The first windows slides from left to right while the second slides from right to left.
41. **Genomic Rapid Algo for String Pm** (GRASPM) - 2009  
 Appeared in [36]  
 Modification of (5). It improve the original algorithm using a filtering method based on an hash function computed on 2-grams in the pattern.
42. **Bounded Boyer-Moore** (BBM) - 2010  
 Appeared in [20]  
 Improvement of (4) using a bounded good-suffix heuristics with constant extra space.

43. **Bounded Fast-Search** (BFS) - 2010  
 Appeared in [20]  
 Improvement of (29) using a bounded good-suffix heuristics with constant extra space.
44. **Bounded Forward-Fast-Search** (BFFS) - 2010  
 Appeared in [20]  
 Improvement of (30) using a bounded forward-good-suffix heuristics with constant extra space.
45. **Fast-Search using Multiple Windows** (FSw) - 2012  
 Appeared in [51]  
 Improvement of (29) using multiple sliding windows. Windows slides from left to right and from right to left.
46. **TVSBS using Multiple Windows** (TVSBSw) - 2012  
 Appeared in [51]  
 Improvement of (34) using multiple sliding windows. Windows slides from left to right and from right to left.
47. **Max Shift Boyer-Moore** (MSBM) - 2012  
 Appeared in [89]  
 Combination of (4) and (5). It takes the maximum shift proposed by the two occurrence heuristics.
48. **Max Shift Horspool** (MSH) - 2012  
 Appeared in [89]  
 Combination of (5) and (4). It takes the maximum shift proposed by the two occurrence heuristics.
49. **Enhanced Two Sliding Windows** (ETSW) - 2012  
 Appeared in [71]  
 Combination of (40) and (27).
50. **Hashq using Multiple Hashing Functions** (MHASHq) - 2012  
 Appeared in [49]  
 Improvement of (38) using multiple hash functions in order to reduce the number false positives.
51. **Enhanced Berry-Ravindran** (RSA) - 2012  
 Appeared in [90]  
 Variant of (27) using four characters to compute the occurrence heuristics.
52. **Enhanced RS-A** (ERSA) - 2013  
 Appeared in [94]  
 Combination of (51) and (40).
53. **Improved Occurrence Heuristics** (IOM) - 2013  
 Appeared in [23,24]  
 Improvement of (5). It uses an improvement of the occurrence heuristics.

54. **Worst Occurrence Heuristics** (WOM) - 2013  
 Appeared in [23,24]  
 Improvement of (5). It computes the occurrence heuristics on the positions which leads to the average maximal shift.
55. **Jumping Occurrence Heuristics** (JOM) - 2013  
 Appeared in [23,24]  
 Improvement of (27). It computes the occurrence heuristics on two non consecutive characters with a distance depending on the average maximal shift.
56. **Simple String Matching** (SSM) - 2015  
 Appeared in [2]  
 Modification of (5). It scans the text from left to right and matches the pattern from right to left.
57. **Quantum Leap Quick-Search** (QLQS) - 2015  
 Appeared in [98]  
 Improvement of (10). It improves the shift performed by the occurrence heuristics by computing the shift to left performed by the reverse of the pattern at a given fixed distance from the current window.
58. **Enhanced ERS-A** (EERSA) - 2015  
 Appeared in [95]  
 Combination of (52) and (49).
59. **Four Sliding Windows** (FSW) - 2015  
 Appeared in [68]  
 Improvement of (52) using four sliding windows.
60. **Skip-Search using  $q$ -grams** (SKIPQ) - 2016  
 Appeared in [43]  
 Combination of (24) and (38). It computes buckets of positions for the fingerprint of each  $q$ -gram in the pattern.



## List of Algorithms based on Automata

Also automata play a very important role in the design of efficient string matching algorithms. For instance, the Deterministic Finite Automaton Matcher [31] was one of the first linear-time solutions, whereas the Backward-DAWG-Matching algorithm [35] reached the optimal  $\mathcal{O}(n \log_\sigma(m)/m)$  lower bound time complexity on the average. Both of them are based on finite automata; in particular, they respectively simulate a deterministic automaton for the language  $\Sigma^*p$  and the deterministic suffix automaton of the reverse of  $p$ .

The efficiency of string matching algorithms depends on the underlying automaton used for recognizing the pattern  $p$  and on the encoding used for simulating it. Here we present the list of all string matching algorithms based on deterministic automata.

61. **Deterministic-Finite-Automaton** (DFA) - no date  
 Appeared in [31]  
 Linear algorithm using a deterministic finite state automaton recognizing all string whose suffix is equal to the pattern. Construction of the automaton can be done in  $\mathcal{O}(m)$ -time.
62. **Reverse-Factor** (RF) - 1992  
 Appeared in [77]  
 Combination of (61) and (4). It uses the suffix automaton of the reverse of the pattern. It works in  $\mathcal{O}(nm)$  worst case time.
63. **Simon** (SIM) - 1994  
 Appeared in [92]  
 Modification of (61). It uses the minimal automaton of the pattern.
64. **Turbo-Reverse-Factor** (TRF) - 1994  
 Appeared in [33]  
 Refinement of (62). It remembers the characters matched in the previous attempt. It's complexity is  $\mathcal{O}(n)$  in time.
65. **Forward-DAWG-Matching** (FDM) - 1994  
 Appeared in [35]  
 It is a linear algorithm using the suffix automaton of the pattern.
66. **Backward-DAWG-Matching** (BDM) - 1994  
 Appeared in [35]  
 Variant of (62). It uses the Directed Acyclic Word Graph of the pattern.
67. **Backward-Oracle-Matching** (BOM) - 1999  
 Appeared in [3]  
 Variant of (62). It uses the Factor Oracle of the pattern. It is the first filtering algorithm using automata.

68. **Double Forward DAWG Matching** (DFDM) - 2000  
 Appeared in [4,5]  
 Modification of (65). It works in linear worst case time using a DAWG of the pattern and a DAWG of the reverse of the pattern.
69. **Wide Window** (WW) - 2005  
 Appeared in [63]  
 Combination of (65) and (62). It uses the suffix automaton of the pattern and the prefix automaton of the reverse of the pattern. It has a linear worst case time complexity.
70. **Linear DAWG Matching** (LDM) - 2005  
 Appeared in [63]  
 Combination of (66) and (61). It uses the suffix automaton of the reverse of the pattern deterministic finite state automaton of the pattern.
71. **Improved Linear DAWG Matching** (ILD1) - 2006  
 Appeared in [79]  
 Improvement of (70).
72. **Improved Linear DAWG Matching 2** (ILD2) - 2006  
 Appeared in [79]  
 Improvement of (70).
73. **Extended Backward Oracle Matching** (EBOM) - 2008  
 Appeared in [46,47]  
 Improvement of (67). It performs transitions of two characters at each step.
74. **Forward Backward Oracle Matching** (FBOM) - 2008  
 Appeared in [46,47]  
 Combination of (73) and (10). It performs transitions using the characters which follows the current window of the text.
75. **Succint Backward DAWG Matching** (SBDM) - 2008  
 Appeared in [57]  
 Variant of (66). It is based on the combination of compressed self-indexes and BDM.
76. **Simplified Extended Backward Oracle Matching** (SEBOM) - 2009  
 Appeared in [40]  
 Simplification of (73).
77. **Simplified Forward Backward Oracle Matching** (SFBOM) - 2009  
 Appeared in [40]  
 Simplification of (74).
78. **Backward SNR DAWG Matching** (BSDM) - 2012  
 Appeared in [50]  
 Modification of (66). It uses the DAWG constructed on the longest substring of the pattern with no repetitions of characters.

- 79. **BSDM using  $q$ -grams and shift-xor** (BSDMQx) - 2016  
 Appeared in [42]  
 Variant of (78). It computes the DAWG assuming a super alphabet on  $q$ -grams of the pattern. The  $q$ -grams are associated to a fingerprint value computed by using a shift-xor hash function.
- 80. **BSDM $_{qx}$  using multiple windows** (BSDMQxw) - 2016  
 Appeared in [42]  
 Improvement of (79) which uses multiple sliding windows.

## List of Algorithms based on Bit Parallelism

Bit-parallelism [8,9] is a technique used for simulating nondeterministic automata. Specifically the bit-parallelism technique takes advantage of the intrinsic parallelism of the bitwise operations inside a computer word, allowing to cut down the number of operations that an algorithm performs by a factor up to  $w$ , where  $w$  is the number of bits in the computer word. However the correspondent encoding requires one bit per pattern symbol, for a total of  $m/w$  computer words. Thus, as long as a pattern fits in a computer word, bit-parallel algorithms are extremely fast, otherwise their performances degrades considerably as  $m/w$  grows.

Here is the list of all string matching algorithms based on bit-parallelism.

- 81. **Shift-Or** (SO) - 1989  
 Appeared in [8,9]  
 Simulates the nondeterministic version of the automaton in (61).
- 82. **Shift-And** (SA) - 1989  
 Appeared in [8,9]  
 Simulates the nondeterministic version of the automaton in (61).
- 83. **Backward-Nondeterministic-DAWG-Matching** (BNDM) - 1998  
 Appeared in [82]  
 Simulates the nondeterministic version of the automaton in (66).
- 84. **BNDM for Long patterns** (BNDML) - 2000  
 Appeared in [83]  
 Modification of (83) using multiple words for simulating the nondeterministic automata of long patterns.
- 85. **Simplified BNDM** (SBNDM) - 2003  
 Appeared in [85,81]  
 Improvement of (83).
- 86. **Two-Way Nondeterministic DAWG Matching** (TNDM) - 2003  
 Appeared in [85]  
 A two-way modification of (83). It scans a pattern suffix forward before normal backward scan.

87. **Long patterns BNDM** (LBNDM) - 2003  
 Appeared in [85]  
 Modification of (83) for long patterns. The automaton is constructed over a superimposed pattern constructed on the  $m/w$  consecutive factors of the inout pattern.
88. **Shift Vector Matching** (SVM) - 2003  
 Appeared in [85]  
 Implementation of (4) using bit parallelism. It remembers characters matched during the last attempt.
89. **BNDM with loop-unrolling** (BNDM2) - 2005  
 Appeared in [64]  
 Improvement of (83) using a blind enrolled loop.
90. **Simplified BNDM with loop-unrolling** (SBNDM2) - 2005  
 Appeared in [64]  
 Improvement of (85) using a blind enrolled loop.
91. **BNDM with Horspool Shift** (BNDMBMH) - 2005  
 Appeared in [64]  
 Combination of (83) and (5)
92. **Horspool with BNDM test** (BMHBNBM) - 2005  
 Appeared in [64]  
 Combination of (83) and (5)
93. **Forward Nondeterministic DAWG Matching** (FNNDM) - 2005  
 Appeared in [64]  
 Simulates the nondeterministic version of the automaton in (65).
94. **Bit parallel Wide Window** (BWW) - 2005  
 Appeared in [63]  
 Simulates the nondeterministic version of the automaton in (69).
95. **Average Optimal Shift-Or** (AOSO) - 2005  
 Appeared in [58]  
 Variant of (81) using a superimposed pattern. It improves the original algorithm by shifting of more than one position during the scan.
96. **Fast Average Optimal Shift-Or** (FAOSO) - 2005  
 Appeared in [58]  
 Practical improvement of (95).
97. **Forward BNDM** (FBNDM) - 2008  
 Appeared in [46,47]  
 Simulates the nondeterministic version of the automaton in (74).
98. **Forward Simplified BNDM** (FSBNDM) - 2008  
 Appeared in [46,47]  
 Combination of (85) and (97).

99. **Bit-Parallel Length Invariant Matcher** (BLIM) - 2008  
 Appeared in [75]  
 Quadratic worst case time algorithm using multiple words.
100. **BNDM with  $q$ -grams** (BNDM $q$ ) - 2009  
 Appeared in [38]  
 Variant of (83) using a super alphabet implemented with  $q$ -grams.
101. **Simplified BNDM with  $q$ -grams** (SBNDM $q$ ) - 2009  
 Appeared in [38]  
 Variant of (85) using a super alphabet implemented with  $q$ -grams.
102. **FNDM with  $q$ -grams** (UFNDM $q$ ) - 2009  
 Appeared in [38]  
 Implementation of (93) using  $q$ -grams.
103. **Small Alphabet Bit-Parallel** (SABP) - 2009  
 Appeared in [100]  
 Bit Parallel algorithm designed for searching string over small alphabets. It is based on a position related character matching table.
104. **BNDM with Extended Shifts** (BXS) - 2010  
 Appeared in [39]  
 Modification of (83) for long patterns. The automaton is constructed over a superimposed pattern constructed on the  $m/w$  consecutive factors of the input pattern.
105. **BNDM $q$  Long** (BQL) - 2010  
 Appeared in [39]  
 Modification of (83) for long patterns. It increases the effective alphabet size by using overlapping  $q$ -grams.
106. **Q-Gram Filtering** (QF) - 2010  
 Appeared in [39]  
 It uses a filtering approach based on consecutive  $q$ -grams in the text.
107. **Bit-Parallel<sup>2</sup> Wide-Window** (BP2WW) - 2010  
 Appeared in [25]  
 Improvement of (69) using two sliding windows inspected in parallel.
108. **Bit-Parallel Wide-Window<sup>2</sup>** (BPWW2) - 2010  
 Appeared in [25]  
 Improvement of (69) using two sliding windows inspected in parallel.
109. **Factorized Shift-And** (KSA) - 2010  
 Appeared in [26,27]  
 Simulates the nondeterministic version of the automaton in (61) using a more compact representation of the automaton. It is based on a factorization of the pattern.

110. **Factorized BNDM** (KBNDM) - 2010  
 Appeared in [26,27]  
 Simulates the nondeterministic version of the automaton in (83) using a more compact representation of the automaton. It is based on a factorization of the pattern.
111. **Forward SBNDM using  $q$ -grams and Lookahead** (FSBNDM<sub>QF</sub>) - 2011  
 Appeared in [86,87]  
 Improvement of (98) which implements a super alphabet with  $q$ -grams and computing the shift with a lookahead of several characters.
112. **SBNDM using Multiple Windows** (SBNDM<sub>w</sub>) - 2012  
 Appeared in [51]  
 Improvement of (85) using multiple sliding windows.
113. **Forward SBNDM using Multiple Windows** (FSBNDM<sub>w</sub>) - 2012  
 Appeared in [51]  
 Improvement of (98) using multiple sliding windows.
114. **Improved Two-Way Shift-And** (TSA) - 2014  
 Appeared in [37]  
 Improvement of (82) using several sliding windows inspected i parallel.
115. **Improved Two-Way Shift-Or** (TSO) - 2014  
 Appeared in [37]  
 Improvement of (81) using several sliding windows inspected i parallel.
116. **Two-Way Shift-And using  $q$ -grams** (TSA<sub>Q</sub>) - 2014  
 Appeared in [37]  
 Improvement of (82) using several sliding windows inspected i parallel. It improves the performances by simulating a super alphabet with  $q$ -grams
117. **Two-Way Shift-Or using  $q$ -grams** (TSO<sub>Q</sub>) - 2014  
 Appeared in [37]  
 Improvement of (81) using several sliding windows inspected i parallel. It improves the performances by simulating a super alphabet with  $q$ -grams

## List of Algorithms based on Packed String Matching

In the packed string matching technique multiple characters are packed into one larger word, so that the characters can be compared in bulk rather than individually. In this context, if the characters of a string are drawn from an alphabet of size  $\sigma$ , then  $\lfloor \frac{w}{\log \sigma} \rfloor$  different characters fit in a single word, using  $\lceil \log \sigma \rceil$  bits per characters. The packing factor is  $\alpha = \lfloor \frac{w}{\log \sigma} \rfloor$ .

118. **Super Alphabet Simulation (SAS)** - 2002  
 Appeared in [56]  
 General approach for packed string matching based on a tabulation technique.
119. **Packed String Search (PSS)** - 2009  
 Appeared in [16,17]  
 Variant of (3) applied to packed string matching.
120. **Streaming SIMD Extensions Filter (SSEF)** - 2009  
 Appeared in [76]  
 Extension of (24) to packed string matching. It is implemented using SSE instructions. It works only for long patterns.
121. **Packed Belazzougui (PB)** - 2010  
 Appeared in [10,11]  
 Efficient packed string matching algorithm which works in  $\mathcal{O}(n/\alpha + occ)$ -time when  $\alpha \leq m \leq n/\alpha$ .
122. **Packed Belazzougui-Raffinot (PBR)** - 2010  
 Appeared in [12]  
 Efficient packed string matching algorithm which works in  $\mathcal{O}(nm)$  worst case time complexity.
123. **Crochemore-Perrin algorithm using SSE instructions (SSECP)** - 2011  
 Appeared in [13,14]  
 Modification of (14) for packed string matching. It is based on two specialized packed string instructions.
124. **Exact Packed String Matching (EPSM)** - 2013  
 Appeared in [44,45]  
 Improvement of (120). It uses SSE instruction to speed up searching. It consists in four different algorithms depending on the length of the pattern.

## References

1. Ahmed, M., Kaykobad, M., Chowdhury, R.A.: A new string matching algorithm. *Int. J. Comput. Math.* 80(7), 825–834 (2003), <http://dx.doi.org/10.1080/0020716031000087113>
2. Al-Ssulami, A.M.: Hybrid string matching algorithm with a pivot. *J. Information Science* 41(1), 82–88 (2015), <http://dx.doi.org/10.1177/0165551514555668>
3. Allauzen, C., Crochemore, M., Raffinot, M.: Factor oracle: A new structure for pattern matching. In: Pavelka, J., Tel, G., Bartosek, M. (eds.) *SOFSEM '99, Theory and Practice of Informatics*, 26th Conference on Current Trends in Theory and Practice of Informatics, Milovy, Czech Republic, November 27 - December 4, 1999, Proceedings. *Lecture Notes in Computer Science*, vol. 1725, pp. 295–310. Springer (1999), [http://dx.doi.org/10.1007/3-540-47849-3\\_18](http://dx.doi.org/10.1007/3-540-47849-3_18)
4. Allauzen, C., Raffinot, M.: Simple optimal string matching algorithm. In: Giancarlo, R., Sankoff, D. (eds.) *Combinatorial Pattern Matching*, 11th Annual Symposium, CPM 2000, Montreal, Canada, June 21-23, 2000, Proceedings. *Lecture Notes in Computer Science*, vol. 1848, pp. 364–374. Springer (2000), [http://dx.doi.org/10.1007/3-540-45123-4\\_30](http://dx.doi.org/10.1007/3-540-45123-4_30)
5. Allauzen, C., Raffinot, M.: Simple optimal string matching algorithm. *J. Algorithms* 36(1), 102–116 (2000), <http://dx.doi.org/10.1006/jagm.2000.1087>
6. Apostolico, A., Crochemore, M.: Optimal canonization of all substrings of a string. *Inf. Comput.* 95(1), 76–95 (1991), [http://dx.doi.org/10.1016/0890-5401\(91\)90016-U](http://dx.doi.org/10.1016/0890-5401(91)90016-U)
7. Apostolico, A., Giancarlo, R.: The boyer-moore-galil string searching strategies revisited. *SIAM J. Comput.* 15(1), 98–105 (1986), <http://dx.doi.org/10.1137/0215007>
8. Baeza-Yates, R.A., Gonnet, G.H.: A new approach to text searching. In: Belkin, N.J., van Rijsbergen, C.J. (eds.) *SIGIR'89, 12th International Conference on Research and Development in Information Retrieval*, Cambridge, Massachusetts, USA, June 25-28, 1989, Proceedings. pp. 168–175. ACM (1989), <http://doi.acm.org/10.1145/75334.75352>
9. Baeza-Yates, R.A., Gonnet, G.H.: A new approach to text searching. *Commun. ACM* 35(10), 74–82 (1992), <http://doi.acm.org/10.1145/135239.135243>
10. Belazzougui, D.: Worst case efficient single and multiple string matching in the RAM model. In: Iliopoulos, C.S., Smyth, W.F. (eds.) *Combinatorial Algorithms - 21st International Workshop, IWOCA 2010, London, UK, July 26-28, 2010, Revised Selected Papers. Lecture Notes in Computer Science*, vol. 6460, pp. 90–102. Springer (2010), [http://dx.doi.org/10.1007/978-3-642-19222-7\\_10](http://dx.doi.org/10.1007/978-3-642-19222-7_10)
11. Belazzougui, D.: Worst-case efficient single and multiple string matching on packed texts in the word-ram model. *J. Discrete Algorithms* 14, 91–106 (2012), <http://dx.doi.org/10.1016/j.jda.2011.12.011>
12. Belazzougui, D., Raffinot, M.: Average optimal string matching in packed strings. In: Spirakis, P.G., Serna, M.J. (eds.) *Algorithms and Complexity*, 8th International Conference, CIAC 2013, Barcelona, Spain, May 22-24, 2013. Proceedings. *Lecture Notes in Computer Science*, vol. 7878, pp. 37–48. Springer (2013), [http://dx.doi.org/10.1007/978-3-642-38233-8\\_4](http://dx.doi.org/10.1007/978-3-642-38233-8_4)
13. Ben-Kiki, O., Bille, P., Breslauer, D., Gasieniec, L., Grossi, R., Weimann, O.: Optimal packed string matching. In: Chakraborty, S., Kumar, A. (eds.) *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2011, December 12-14, 2011, Mumbai, India. LIPIcs*,



- vol. 13, pp. 423–432. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik (2011), <http://dx.doi.org/10.4230/LIPIcs.FSTTCS.2011.423>
14. Ben-Kiki, O., Bille, P., Breslauer, D., Gasieniec, L., Grossi, R., Weimann, O.: Towards optimal packed string matching. *Theor. Comput. Sci.* 525, 111–129 (2014), <http://dx.doi.org/10.1016/j.tcs.2013.06.013>
  15. Berry, T., Ravindran, S.: A fast string matching algorithm and experimental results. In: Holub, J., Simánek, M. (eds.) *Proceedings of the Prague Stringology Club Workshop 1999*, Prague, Czech Republic, July 8–9, 1999. pp. 16–28. Department of Computer Science and Engineering, Faculty of Electrical Engineering, Czech Technical University (1999), <http://www.stringology.org/event/1999/p2.html>
  16. Bille, P.: Fast searching in packed strings. In: Kucherov, G., Ukkonen, E. (eds.) *Combinatorial Pattern Matching, 20th Annual Symposium, CPM 2009*, Lille, France, June 22–24, 2009, *Proceedings. Lecture Notes in Computer Science*, vol. 5577, pp. 116–126. Springer (2009), [http://dx.doi.org/10.1007/978-3-642-02441-2\\_11](http://dx.doi.org/10.1007/978-3-642-02441-2_11)
  17. Bille, P.: Fast searching in packed strings. *J. Discrete Algorithms* 9(1), 49–56 (2011), <http://dx.doi.org/10.1016/j.jda.2010.09.003>
  18. Boyer, R.S., Moore, J.S.: A fast string searching algorithm. *Commun. ACM* 20(10), 762–772 (1977), <http://doi.acm.org/10.1145/359842.359859>
  19. Cantone, D., Faro, S.: Searching for a substring with constant extra-space complexity. In: Ferragina, P., Grossi, R. (eds.) *Proc. of Third International Conference on Fun with algorithms*. pp. 118–131 (2004)
  20. Cantone, D., Cristofaro, S., Faro, S.: A space-efficient implementation of the good-suffix heuristic. In: Holub, J., Zdárek, J. (eds.) *Proceedings of the Prague Stringology Conference 2010*, Prague, Czech Republic, August 30 - September 1, 2010. pp. 63–75. Prague Stringology Club, Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague (2010), <http://www.stringology.org/event/2010/p06.html>
  21. Cantone, D., Faro, S.: Fast-search: A new efficient variant of the boyer-moore string matching algorithm. In: Jansen, K., Margraf, M., Mastrolilli, M., Rolim, J.D.P. (eds.) *Experimental and Efficient Algorithms, Second International Workshop, WEA 2003*, Ascona, Switzerland, May 26–28, 2003, *Proceedings. Lecture Notes in Computer Science*, vol. 2647, pp. 47–58. Springer (2003), [http://dx.doi.org/10.1007/3-540-44867-5\\_4](http://dx.doi.org/10.1007/3-540-44867-5_4)
  22. Cantone, D., Faro, S.: Fast-search algorithms: New efficient variants of the boyer-moore pattern-matching algorithm. *Journal of Automata, Languages and Combinatorics* 10(5/6), 589–608 (2005)
  23. Cantone, D., Faro, S.: Improved and self-tuned occurrence heuristics. In: Holub, J., Zdárek, J. (eds.) *Proceedings of the Prague Stringology Conference 2013*, Prague, Czech Republic, September 2–4, 2013. pp. 92–106. Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague (2013), <http://www.stringology.org/event/2013/p09.html>
  24. Cantone, D., Faro, S.: Improved and self-tuned occurrence heuristics. *J. Discrete Algorithms* 28, 73–84 (2014), <http://dx.doi.org/10.1016/j.jda.2014.07.006>
  25. Cantone, D., Faro, S., Giaquinta, E.: Bit-(parallelism)<sup>2</sup>: Getting to the next level of parallelism. In: Boldi, P., Gargano, L. (eds.) *Fun with Algorithms, 5th International Conference, FUN 2010*, Ischia, Italy, June 2–4, 2010. *Proceedings*.

- Lecture Notes in Computer Science, vol. 6099, pp. 166–177. Springer (2010), [http://dx.doi.org/10.1007/978-3-642-13122-6\\_18](http://dx.doi.org/10.1007/978-3-642-13122-6_18)
26. Cantone, D., Faro, S., Giaquinta, E.: A compact representation of nondeterministic (suffix) automata for the bit-parallel approach. In: Amir, A., Parida, L. (eds.) *Combinatorial Pattern Matching, 21st Annual Symposium, CPM 2010*, New York, NY, USA, June 21–23, 2010. *Proceedings. Lecture Notes in Computer Science*, vol. 6129, pp. 288–298. Springer (2010), [http://dx.doi.org/10.1007/978-3-642-13509-5\\_26](http://dx.doi.org/10.1007/978-3-642-13509-5_26)
  27. Cantone, D., Faro, S., Giaquinta, E.: A compact representation of nondeterministic (suffix) automata for the bit-parallel approach. *Inf. Comput.* 213, 3–12 (2012), <http://dx.doi.org/10.1016/j.ic.2011.03.006>
  28. Charras, C., Lecroq, T., Pehoushek, J.D.: A very fast string matching algorithm for small alphabets and long patterns (extended abstract). In: *Farach-Colton [41]*, pp. 55–64, <http://dx.doi.org/10.1007/BFb0030780>
  29. Colussi, L.: Correctness and efficiency of the pattern matching algorithms. *Inf. Comput.* 95(2), 225–251 (1991), [http://dx.doi.org/10.1016/0890-5401\(91\)90046-5](http://dx.doi.org/10.1016/0890-5401(91)90046-5)
  30. Colussi, L.: Fastest pattern matching in strings. *J. Algorithms* 16(2), 163–189 (1994), <http://dx.doi.org/10.1006/jagm.1994.1008>
  31. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: *Introduction to Algorithms* (3. ed.). MIT Press (2009), <http://mitpress.mit.edu/books/introduction-algorithms>
  32. Crochemore, M.: String-matching on ordered alphabets. *Theor. Comput. Sci.* 92(1), 33–47 (1992), [http://dx.doi.org/10.1016/0304-3975\(92\)90134-2](http://dx.doi.org/10.1016/0304-3975(92)90134-2)
  33. Crochemore, M., Czumaj, A., Gasieniec, L., Jarominek, S., Lecroq, T., Plandowski, W., Rytter, W.: Speeding up two string-matching algorithms. *Algorithmica* 12(4/5), 247–267 (1994), <http://dx.doi.org/10.1007/BF01185427>
  34. Crochemore, M., Perrin, D.: Two-way string matching. *J. ACM* 38(3), 651–675 (1991), <http://doi.acm.org/10.1145/116825.116845>
  35. Crochemore, M., Rytter, W.: *Text Algorithms*. Oxford University Press (1994), <http://www-igm.univ-mlv.fr/~mac/REC/B1.html>
  36. Deusdado, S.A.D., Carvalho, P.: Grapm: an efficient algorithm for exact pattern-matching in genomic sequences. *IJBRA* 5(4), 385–401 (2009), <http://dx.doi.org/10.1504/IJBRA.2009.027510>
  37. Durian, B., Chhabra, T., Ghuman, S.S., Hirvola, T., Peltola, H., Tarhio, J.: Improved two-way bit-parallel search. In: Holub, J., Zdárek, J. (eds.) *Proceedings of the Prague Stringology Conference 2014*, Prague, Czech Republic, September 1–3, 2014. pp. 71–83. Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague (2014), <http://www.stringology.org/event/2014/p08.html>
  38. Durian, B., Holub, J., Peltola, H., Tarhio, J.: Tuning BNDM with q-grams. In: Finocchi, I., Hershberger, J. (eds.) *Proceedings of the Eleventh Workshop on Algorithm Engineering and Experiments, ALENEX 2009*, New York, New York, USA, January 3, 2009. pp. 29–37. SIAM (2009), <http://dx.doi.org/10.1137/1.9781611972894.3>
  39. Durian, B., Peltola, H., Salmela, L., Tarhio, J.: Bit-parallel search algorithms for long patterns. In: Festa, P. (ed.) *Experimental Algorithms, 9th International Symposium, SEA 2010*, Ischia Island, Naples, Italy, May 20–22, 2010. *Proceedings. Lecture Notes in Computer Science*, vol. 6049, pp. 129–140. Springer (2010), [http://dx.doi.org/10.1007/978-3-642-13193-6\\_12](http://dx.doi.org/10.1007/978-3-642-13193-6_12)

40. Fan, H., Yao, N., Ma, H.: Fast variants of the backward-oracle-matching algorithm. In: Proceedings of the 2009 Fourth International Conference on Internet Computing for Science and Engineering. pp. 56–59. ICICSE '09, IEEE Computer Society, Washington, DC, USA (2009), <http://dx.doi.org/10.1109/ICICSE.2009.53>
41. Farach-Colton, M. (ed.): Combinatorial Pattern Matching, 9th Annual Symposium, CPM 98, Piscataway, New Jersey, USA, July 20–22, 1998, Proceedings, Lecture Notes in Computer Science, vol. 1448. Springer (1998)
42. Faro, S.: Evaluation and improvement of fast algorithms for exact matching on genome sequences. In: 3rd International Conference on Algorithms for Computational Biology. Lecture Notes in Computer Science, Springer (2016)
43. Faro, S.: A very fast string matching algorithm based on condensed alphabets. In: Algorithmic Aspects in Information and Management - 10th International Conference, AAIM 2016. Proceedings. Lecture Notes in Computer Science, Springer (2016)
44. Faro, S., Külekci, M.O.: Fast packed string matching for short patterns. In: Sanders, P., Zeh, N. (eds.) Proceedings of the 15th Meeting on Algorithm Engineering and Experiments, ALENEX 2013, New Orleans, Louisiana, USA, January 7, 2013. pp. 113–121. SIAM (2013), <http://dx.doi.org/10.1137/1.9781611972931.10>
45. Faro, S., Külekci, M.O.: Fast and flexible packed string matching. J. Discrete Algorithms 28, 61–72 (2014), <http://dx.doi.org/10.1016/j.jda.2014.07.003>
46. Faro, S., Lecroq, T.: Efficient variants of the backward-oracle-matching algorithm. In: Holub, J., Zdárek, J. (eds.) Proceedings of the Prague Stringology Conference 2008, Prague, Czech Republic, September 1–3, 2008. pp. 146–160. Prague Stringology Club, Department of Computer Science and Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague (2008), <http://www.stringology.org/event/2008/p14.html>
47. Faro, S., Lecroq, T.: Efficient variants of the backward-oracle-matching algorithm. Int. J. Found. Comput. Sci. 20(6), 967–984 (2009), <http://dx.doi.org/10.1142/S0129054109006991>
48. Faro, S., Lecroq, T.: The exact string matching problem: a comprehensive experimental evaluation. CoRR abs/1012.2547 (2010), <http://arxiv.org/abs/1012.2547>
49. Faro, S., Lecroq, T.: Fast searching in biological sequences using multiple hash functions. In: 12th IEEE International Conference on Bioinformatics & Bioengineering, BIBE 2012, Larnaca, Cyprus, November 11–13, 2012. pp. 175–180. IEEE Computer Society (2012), <http://dx.doi.org/10.1109/BIBE.2012.6399669>
50. Faro, S., Lecroq, T.: A fast suffix automata based algorithm for exact online string matching. In: Moreira, N., Reis, R. (eds.) Implementation and Application of Automata - 17th International Conference, CIAA 2012, Porto, Portugal, July 17–20, 2012. Proceedings. Lecture Notes in Computer Science, vol. 7381, pp. 149–158. Springer (2012), [http://dx.doi.org/10.1007/978-3-642-31606-7\\_13](http://dx.doi.org/10.1007/978-3-642-31606-7_13)
51. Faro, S., Lecroq, T.: A multiple sliding windows approach to speed up string matching algorithms. In: Klasing, R. (ed.) Experimental Algorithms - 11th International Symposium, SEA 2012, Bordeaux, France, June 7–9, 2012. Proceedings. Lecture Notes in Computer Science, vol. 7276, pp. 172–183. Springer (2012), [http://dx.doi.org/10.1007/978-3-642-30850-5\\_16](http://dx.doi.org/10.1007/978-3-642-30850-5_16)
52. Faro, S., Lecroq, T.: The exact online string matching problem: A review of the most recent results. ACM Comput. Surv. 45(2), 13 (2013), <http://doi.acm.org/10.1145/2431211.2431212>

53. Finkel, A., Jantzen, M. (eds.): STACS 92, 9th Annual Symposium on Theoretical Aspects of Computer Science, Cachan, France, February 13-15, 1992, Proceedings, Lecture Notes in Computer Science, vol. 577. Springer (1992)
54. Franek, F., Jennings, C.G., Smyth, W.F.: A simple fast hybrid pattern-matching algorithm. In: Apostolico, A., Crochemore, M., Park, K. (eds.) Combinatorial Pattern Matching, 16th Annual Symposium, CPM 2005, Jeju Island, Korea, June 19-22, 2005, Proceedings. Lecture Notes in Computer Science, vol. 3537, pp. 288–297. Springer (2005), [http://dx.doi.org/10.1007/11496656\\_25](http://dx.doi.org/10.1007/11496656_25)
55. Franek, F., Jennings, C.G., Smyth, W.F.: A simple fast hybrid pattern-matching algorithm. *J. Discrete Algorithms* 5(4), 682–695 (2007), <http://dx.doi.org/10.1016/j.jda.2006.11.004>
56. Fredriksson, K.: Faster string matching with super-alphabets. In: Laender, A.H.F., Oliveira, A.L. (eds.) String Processing and Information Retrieval, 9th International Symposium, SPIRE 2002, Lisbon, Portugal, September 11-13, 2002, Proceedings. Lecture Notes in Computer Science, vol. 2476, pp. 44–57. Springer (2002), [http://dx.doi.org/10.1007/3-540-45735-6\\_5](http://dx.doi.org/10.1007/3-540-45735-6_5)
57. Fredriksson, K.: Succinct backward-dawg-matching. *ACM Journal of Experimental Algorithmics* 13 (2008), <http://doi.acm.org/10.1145/1412228.1455263>
58. Fredriksson, K., Grabowski, S.: Practical and optimal string matching. In: Consens, M.P., Navarro, G. (eds.) String Processing and Information Retrieval, 12th International Conference, SPIRE 2005, Buenos Aires, Argentina, November 2-4, 2005, Proceedings. Lecture Notes in Computer Science, vol. 3772, pp. 376–387. Springer (2005), [http://dx.doi.org/10.1007/11575832\\_42](http://dx.doi.org/10.1007/11575832_42)
59. Galil, Z., Giancarlo, R.: On the exact complexity of string matching: Upper bounds. *SIAM J. Comput.* 21(3), 407–437 (1992), <http://dx.doi.org/10.1137/0221028>
60. Galil, Z., Seiferas, J.I.: Time-space-optimal string matching. In: Proceedings of the 13th Annual ACM Symposium on Theory of Computing, May 11-13, 1981, Milwaukee, Wisconsin, USA. pp. 106–113. ACM (1981), <http://doi.acm.org/10.1145/800076.802463>
61. Galil, Z., Seiferas, J.I.: Time-space-optimal string matching. *J. Comput. Syst. Sci.* 26(3), 280–294 (1983), [http://dx.doi.org/10.1016/0022-0000\(83\)90002-8](http://dx.doi.org/10.1016/0022-0000(83)90002-8)
62. Hancart, C.: Analyse exacte et en moyenne d’algorithmes de recherche d’un motif dans un texte. Thèse de doctorat, Université Paris 7 (1993)
63. He, L., Fang, B., Sui, J.: The wide window string matching algorithm. *Theor. Comput. Sci.* 332(1-3), 391–404 (2005), <http://dx.doi.org/10.1016/j.tcs.2004.12.002>
64. Holub, J., Durian, B.: Talk: Fast variants of bit parallel approach to suffix automata. In: The Second Haifa Annual International Stringology Research Workshop of the Israeli Science Foundation, <http://www.cri.haifa.ac.il/events/2005/string/presentations/Holub.pdf> (2005)
65. Horspool, R.N.: Practical fast searching in strings. *Softw., Pract. Exper.* 10(6), 501–506 (1980), <http://dx.doi.org/10.1002/spe.4380100608>
66. Huang, Y., Ping, L., Pan, X., Cai, G.: A fast exact pattern matching algorithm for biological sequences. In: Proceedings of the 2008 International Conference on BioMedical Engineering and Informatics, BMEI 2008, May 28-30, 2008, Sanya, Hainan, China - Volume 1. pp. 8–12. IEEE Computer Society (2008), <http://dx.doi.org/10.1109/BMEI.2008.154>
67. Hudaib, A., Al-Khalid, R., Suleiman, D., Itriq, M., Al-Anani, A.: A fast pattern matching algorithm with two sliding windows (tsw). *Journal of Computer Science* 4(5), 393–401 (2008)

68. Hudaib, A., Al-Khalid, R., Al-Anani, A., Itriq, M., Suleiman, D.: Four sliding windows pattern matching algorithm (fsw). *Journal of Software Engineering and Applications* 8, 154–165 (2015)
69. Hume, A., Sunday, D.: Fast string searching. In: *Proceedings of the Summer 1991 USENIX Conference*, Nashville, TE, USA, June 1991. pp. 221–234. USENIX Association (1991)
70. Hume, A., Sunday, D.: Fast string searching. *Softw., Pract. Exper.* 21(11), 1221–1248 (1991), <http://dx.doi.org/10.1002/spe.4380211105>
71. Itriq, M., Hudaib, A., Al-Anani, A., Al-Khalid, R., Suleiman, D.: Enhanced two sliding windows algorithm for pattern matching (etsw). *Journal of American Science* 8, 60–616 (2012)
72. Kalsi, P., Peltola, H., Tarhio, J.: Comparison of exact string matching algorithms for biological sequences. In: Elloumi, M., Küng, J., Linial, M., Murphy, R.F., Schneider, K., Toma, C. (eds.) *Bioinformatics Research and Development, Second International Conference, BIRD 2008*, Vienna, Austria, July 7-9, 2008, *Proceedings. Communications in Computer and Information Science*, vol. 13, pp. 417–426. Springer (2008), [http://dx.doi.org/10.1007/978-3-540-70600-7\\_31](http://dx.doi.org/10.1007/978-3-540-70600-7_31)
73. Karp, R.M., Rabin, M.O.: Efficient randomized pattern-matching algorithms. *IBM Journal of Research and Development* 31(2), 249–260 (1987), <http://dx.doi.org/10.1147/rd.312.0249>
74. Knuth, D.E., Jr., J.H.M., Pratt, V.R.: Fast pattern matching in strings. *SIAM J. Comput.* 6(2), 323–350 (1977), <http://dx.doi.org/10.1137/0206024>
75. Külekci, M.O.: A method to overcome computer word size limitation in bit-parallel pattern matching. In: Hong, S., Nagamochi, H., Fukunaga, T. (eds.) *Algorithms and Computation, 19th International Symposium, ISAAC 2008*, Gold Coast, Australia, December 15-17, 2008. *Proceedings. Lecture Notes in Computer Science*, vol. 5369, pp. 496–506. Springer (2008), [http://dx.doi.org/10.1007/978-3-540-92182-0\\_45](http://dx.doi.org/10.1007/978-3-540-92182-0_45)
76. Külekci, M.O.: Filter based fast matching of long patterns by using SIMD instructions. In: Holub, J., Zdárek, J. (eds.) *Proceedings of the Prague Stringology Conference 2009*, Prague, Czech Republic, August 31 - September 2, 2009. pp. 118–128. Prague Stringology Club, Department of Computer Science and Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague (2009), <http://www.stringology.org/event/2009/p11.html>
77. Lecroq, T.: A variation on the boyer-moore algorithm. *Theor. Comput. Sci.* 92(1), 119–144 (1992), [http://dx.doi.org/10.1016/0304-3975\(92\)90139-7](http://dx.doi.org/10.1016/0304-3975(92)90139-7)
78. Lecroq, T.: Fast exact string matching algorithms. *Inf. Process. Lett.* 102(6), 229–235 (2007), <http://dx.doi.org/10.1016/j.ipl.2007.01.002>
79. Liu, C., Wang, Y., Liu, D., Li, D.: Two improved single pattern matching algorithms. In: *16th International Conference on Artificial Reality and Telexistence, ICAT 2006*, Hangzhou, China, November 29 - December 1, 2006, *Workshops Proceedings*. pp. 419–422. IEEE Computer Society (2006), <http://dx.doi.org/10.1109/ICAT.2006.134>
80. Morris, Jr, J.H., Pratt, V.R.: A linear pattern-matching algorithm. Report 40, University of California, Berkeley (1970)
81. Navarro, G.: Nr-grep: a fast and flexible pattern-matching tool. *Softw., Pract. Exper.* 31(13), 1265–1312 (2001), <http://dx.doi.org/10.1002/spe.411>
82. Navarro, G., Raffinot, M.: A bit-parallel approach to suffix automata: Fast extended string matching. In: *Farach-Colton [41]*, pp. 14–33, <http://dx.doi.org/10.1007/BFb0030778>

83. Navarro, G., Raffinot, M.: Fast and flexible string matching by combining bit-parallelism and suffix automata. *ACM Journal of Experimental Algorithmics* 5, 4 (2000), <http://doi.acm.org/10.1145/351827.384246>
84. Nebel, M.E.: Fast string matching by using probabilities: On an optimal mismatch variant of horspool's algorithm. *Theor. Comput. Sci.* 359(1-3), 329–343 (2006), <http://dx.doi.org/10.1016/j.tcs.2006.05.028>
85. Peltola, H., Tarhio, J.: Alternative algorithms for bit-parallel string matching. In: Nascimento, M.A., de Moura, E.S., Oliveira, A.L. (eds.) *String Processing and Information Retrieval*, 10th International Symposium, SPIRE 2003, Manaus, Brazil, October 8-10, 2003, Proceedings. *Lecture Notes in Computer Science*, vol. 2857, pp. 80–94. Springer (2003), [http://dx.doi.org/10.1007/978-3-540-39984-1\\_7](http://dx.doi.org/10.1007/978-3-540-39984-1_7)
86. Peltola, H., Tarhio, J.: Variations of forward-sbndm. In: Holub, J., Zdárek, J. (eds.) *Proceedings of the Prague Stringology Conference 2011*, Prague, Czech Republic, August 29-31, 2011. pp. 3–14. Prague Stringology Club, Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague (2011), <http://www.stringology.org/event/2011/p02.html>
87. Peltola, H., Tarhio, J.: String matching with lookahead. *Discrete Applied Mathematics* 163, 352–360 (2014), <http://dx.doi.org/10.1016/j.dam.2013.10.034>
88. Raita, T.: Tuning the boyer-moore-horspool string searching algorithm. *Softw., Pract. Exper.* 22(10), 879–884 (1992), <http://dx.doi.org/10.1002/spe.4380221006>
89. Sahli, M., Shibuya, T.: Max-shift BM and max-shift horspool: Practical fast exact string matching algorithms. *JIP* 20(2), 419–425 (2012), <http://dx.doi.org/10.2197/ipsjjip.20.419>
90. Senapati, K., Mal, S., Sahoo, G.: Rs-a fast pattern matching algorithm for biological sequences. *International Journal of Engineering and Innovative Technology* 1, 116–118 (2012)
91. Sheik, S.S., Aggarwal, S.K., Poddar, A., Balakrishnan, N., Sekar, K.: A FAST pattern matching algorithm. *Journal of Chemical Information and Modeling* 44(4), 1251–1256 (2004), <http://dx.doi.org/10.1021/ci030463z>
92. Simon, I.: String matching algorithms and automata. In: Karhumäki, J., Maurer, H.A., Rozenberg, G. (eds.) *Results and Trends in Theoretical Computer Science*, Colloquium in Honor of Arto Salomaa, Graz, Austria, June 10-11, 1994, Proceedings. *Lecture Notes in Computer Science*, vol. 812, pp. 386–395. Springer (1994), [http://dx.doi.org/10.1007/3-540-58131-6\\_61](http://dx.doi.org/10.1007/3-540-58131-6_61)
93. Smith, P.D.: Experiments with a very fast substring search algorithm. *Softw., Pract. Exper.* 21(10), 1065–1074 (1991), <http://dx.doi.org/10.1002/spe.4380211006>
94. Suleiman, D., Hudaib, A., Al-Anani, A., R., A.K., Itriq, M.: Ers-a algorithm for pattern matching. *Middle East Journal of Scientific Research* 15, 1067–1075 (2013)
95. Suleiman, D., Itriq, M., Al-Anani, A., Al-Khalid, R., Hudaib, A.: Enhancing ers-a algorithm for pattern matching (eers-a). *Journal of Software Engineering and Applications* 8, 143–153 (2015)
96. Sunday, D.: A very fast substring search algorithm. *Commun. ACM* 33(8), 132–142 (1990), <http://doi.acm.org/10.1145/79173.79184>
97. Sustik, M., Moore, J.: String searching over small alphabets. In: Technical Report TR-07-62. Department of Computer Sciences, University of Texas at Austin (2007)

98. Watson, B.W., Kourie, D.G., Cleophas, L.G.: Quantum leap pattern matching. In: Holub, J., Zdárek, J. (eds.) *Proceedings of the Prague Stringology Conference 2015*, Prague, Czech Republic, August 24-26, 2015. pp. 104–117. Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague (2015), <http://www.stringology.org/event/2015/p11.html>
99. Yao, A.C.: The complexity of pattern matching for a random string. *SIAM J. Comput.* 8(3), 368–387 (1979), <http://dx.doi.org/10.1137/0208029>
100. Zhang, G., Zhu, E., Mao, L., Yin, M.: A bit-parallel exact string matching algorithm for small alphabet. In: Deng, X., Hopcroft, J.E., Xue, J. (eds.) *Frontiers in Algorithmics, Third International Workshop, FAW 2009*, Hefei, China, June 20-23, 2009. *Proceedings. Lecture Notes in Computer Science*, vol. 5598, pp. 336–345. Springer (2009), [http://dx.doi.org/10.1007/978-3-642-02270-8\\_34](http://dx.doi.org/10.1007/978-3-642-02270-8_34)
101. Zhu, R.F., Takaoka, T.: On improving the average case of the Boyer-Moore string matching algorithm. *J. Inform. Process.* 10(3), 173–177 (1987)